Expected-Value Norms on Matrices

Eric C. Howe*
Department of Economics
University of Saskatchewan
Saskatoon, Saskatchewan, Canada S7N 0W0

and

Charles R. Johnson†
Department of Mathematics
College of William and Mary
Williamsburg, Virginia 23185

Submitted by Richard A. Brualdi

ABSTRACT

For any vector norm, the function that assigns to a matrix A the "average" norm of Ax is a generalized matrix norm. Certain properties of such expected-value norms on matrices are noted, as well as a motivating example for error analysis in input-output models. For the l_1 vector norm an explicit formula is given with respect to averaging over the Euclidean unit ball.

INTRODUCTION

Induced matrix norms are a very commonly used type of norm on matrices. Let M_k be the set of complex k-by-k matrices, let $M_{k,l}$ be the set of complex k-by-l matrices, and let C^k be the set of all complex k-dimen-

*The work of this author was supported by the Social Sciences and Humanities Research Council of Canada.

†The work of this author was supported by National Science Foundation grant DMS-8802836 and a grant from the National Security Agency.

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

sional vectors. If $A \in M_k$ and $\|\cdot\|$ is a vector norm on C^k, then the matrix norm induced by $\|\cdot\|$ is defined by

$$|||A||| = \max_{\|x\|=1} \|Ax\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Induced matrix norms have a number of nice properties. One is that they are known to be *minimal* among matrix norms; i.e., for any matrix norm $N(\cdot)$ there exists an induced matrix norm $|||\cdot|||$ such that $|||A||| \leqslant N(A)$ for all $A \in M_n$; and for any induced matrix norm $|||\cdot|||$ and any matrix norm $N(\cdot)$, we have $N(A) \leqslant |||A|||$ for all $A \in M_k$ if and only if $N(A) = |||A|||$ for all $A \in M_k$. [Horn and Johnson (1985, Chapter 5) contains an extensive discussion of norms on matrices, and concepts about norms not defined here may be found there.]

In spite of the above attractive properties, induced norms may make matrices seem to be large in the following sense. The value of $|||A|||$ records the "worst case" (extreme) value by which A stretches a vector in the norm $\|\cdot\|$, and this may be very large relative to the magnification of most vectors by A. In applications involving error analysis, minimization of $|||A|||$ may be misleading in comparison with minimizing a more representative value of $\|Ax\|$. Although all norms on a finite-dimensional vector space are comparable, our purpose here is to introduce another class of norms on matrices derived from norms on vectors. This class focuses upon an "average" value of $\|Ax\|$, and we make several observations about it.

EXPECTED-VALUE GENERALIZED MATRIX NORMS

Let $n_1(\cdot)$ and $n_2(\cdot)$ be two vector norms on C^k. Let z be a random variable on C^k with a uniform distribution on the surface of the unit ball $B(n_2)$ of n_2. The *expected-value norm* on M_k relative to n_1 and n_2 is then defined by

$$G(A; n_1, n_2) = \frac{E[n_1(Az)]}{E[n_1(z)]}.$$

We will usually write $G(A)$ instead of $G(A; n_1, n_2)$. We note that if n_1 and n_2 are the same, then $G(A) = E[n_1(Az)]$.

It is clear that the function G is well defined, and we first note that G is a generalized matrix norm on M_k. Three properties are definitionally required:

PROPERTY 1 (Nonnegativity). $G(A) \geqslant 0$ for all $A \in M_k$, with equality if and only if $A = 0$.

PROPERTY 2 (Homogeneity). $G(\alpha A) = |\alpha|\ G(A)$ for all $\alpha \in C$ and $A \in M_k$.

PROPERTY 3 (Triangle inequality). $G(A+B) \leqslant G(A) + G(B)$ for all $A, B \in M_k$.

Clearly $G(A)$ satisfies nonnegativity and homogeneity. It inherits the triangle inequality from n_1:

$$G(A+B) = \frac{E[n_1((A+B)z)]}{E[n_1(z)]} \leqslant \frac{E[n_1(Az) + n_1(Bz)]}{E[n_1(z)]} = G(A) + G(B).$$

In order to be a matrix norm, $G(A)$ would have to enjoy one additional property:

PROPERTY 4 (Submultiplicativity). $G(AB) \leqslant G(A)G(B)$ for all $A, B \in M_k$.

The question of whether there exist n_1 and n_2 such that $G(A)$ is submultiplicative is addressed in the following observation.

THEOREM 1. *For no pair of vector norms n_1, n_2 on C^k is $G(A; n_1, n_2)$ submultiplicative.*

Proof. Let $|||\cdot|||$ be the matrix norm induced by n_1. Note that

$$G(A) = \frac{E[n_1(Az)]}{E[n_1(z)]} = \frac{E[n_1(z)\{n_1(Az)/n_1(z)\}]}{E[n_1(z)]}$$

$$\leqslant \frac{E[n_1(z)\,|||A|||]}{E[n_1(z)]} = |||A|||$$

However, as noted in the introduction, no matrix norm is less than or equal to an induced norm (other than the induced norm itself). But it is easy to see that for any n_1, n_2, G is not identically equal to $|||\cdot|||$; so $G(A)$ is not submultiplicative. ∎

In fact, we conjecture that no expected-value norm G can be spectrally dominant [the spectral radius of A is $\leqslant G(A)$ for all $A \in M_k$], and this is easily proven when n_1 is monotone.

We note that $G(A) = |||A|||$ can occur for some matrices. In particular, it is clear from the calculation in the proof of Theorem 1 that $G(A) = |||A|||$ if and only if A is a multiple of an isometry for n_1.

The fact that $G(A)$ is not submultiplicative might limit the usefulness of expected-value norms, except that they have the following closely related property.

THEOREM 2. *Let n_1 and n_2 be vector norms on C^k, and let $|||\cdot|||$ be the matrix norm induced by n_1. Then*

$$G(AB; n_1, n_2) \leqslant |||A|||\, G(B; n_1, n_2).$$

Proof. Observe that

$$G(AB) = \frac{E[n_1(ABz)]}{E[n_1(z)]} \leqslant \frac{E[\,|||A|||\, n_1(Bz)]}{E[n_1(z)]} = |||A|||\, G(B). \qquad \blacksquare$$

The inequality in Theorem 2 is tight whenever A is a multiple of an isometry for n_1.

In some applications, the property given in the above theorem is as useful as submultiplicativity. For example, consider $A \in M_n$ with $|||A||| < 1$. If we wanted a bound on $|||A + A^2 + A^3 + \cdots|||$, submultiplicativity (and the triangle inequality) would yield the following:

$$|||A + A^2 + A^3 + \cdots||| \leqslant \frac{|||A|||}{1 - |||A|||}.$$

Similarly, the above theorem guarantees that

$$G(A + A^2 + A^3 + \cdots) \leqslant \frac{G(A)}{1 - |||A|||}.$$

If we let $A = \alpha I$ for $\alpha \in C$, $|\alpha| < 1$, then both sides of the above inequality have values of $\alpha/(1-\alpha)$, so that the bound is tight.

THE EXPECTED-VALUE NORM RESULTING FROM ONE CHOICE OF n_1 AND n_2

The value of $G(A)$ could always be approximated for a given A by applying an integration algorithm to $E[n_1(Az)]$ and $E[n_1(z)]$. However, we give one example of n_1 and n_2 that allows $G(A)$ to be expressed in closed form.

We will let $\|x\|_2 = (\Sigma|x_i|^2)^{0.5}$ denote the Euclidean vector norm and $\|x\|_1 = \Sigma|x_i|$ denote the absolute-sum vector norm. We require two preliminary lemmas. The first is well known.

LEMMA 1. *Let $x \in C^k$ with $\|x\|_2 = 1$, and let $e^{(1)} \in C^k$ be the first vector in the standard basis of C^k. There exists a unitary matrix $U \in M_k$ such that $Ux = e^{(1)}$.*

LEMMA 2. *Let z be a k-dimensional random variable with a uniform distribution on $B(\|\cdot\|_2)$. Let $\xi \in R$ be given by*

$$\xi = \frac{1}{\int_{B(\|\cdot\|_2)} 1\, dz}.$$

Let $x \in C^k$. Then

$$E[|x^*z|] = \int_{B(\|\cdot\|_2)} \xi|x^*z|\, dz = \|x\|_2 \int_{B(\|\cdot\|_2)} \xi|z_1|\, dz = \|x\|_2 E[|z_1|],$$

in which z_1 is the first coordinate of z.

Proof. Let $y = x/\|x\|_2$. The proof follows from a change of variable. Let U be a unitary matrix given by Lemma 1 for the vector y. Let $r = Uz$, so that $z = U^*r$. Note that the Jacobian of this transformation equals 1, since U is unitary; note also that the region of integration $B(\|\cdot\|_2)$ is unaffected by the change of variable, since unitary matrices are in the isometry group of the Euclidean norm. Thus

$$\begin{aligned}
E[|x^*z|] &= \int_{B(\|\cdot\|_2)} \xi |x^*z|\, dz \\
&= \int_{B(\|\cdot\|_2)} \xi |x^*U^*r|\, dr \\
&= \|x\|_2 \int_{B(\|\cdot\|_2)} \xi |(Uy)^*r|\, dr \\
&= \|x\|_2 \int_{B(\|\cdot\|_2)} \xi |r_1|\, dr \\
&= \|x\|_2 \int_{B(\|\cdot\|_2)} \xi |z_1|\, dz \\
&= \|x\|_2 E[|z_1|]. \qquad \blacksquare
\end{aligned}$$

We note that $E[|z_1|]$ is just a (complicated) constant, depending upon k, and that $E[|z_i|] = E[|z_1|]$, $i = 1,\dots,k$. We are not interested in its precise value, as it cancels out in the following calculations.

THEOREM 3. *Let $A \in M_k$, and partition A by rows, so that $A^{\mathsf{T}} = [A_1 \cdots A_k]$. Then*

$$G(A; \|\cdot\|_1, \|\cdot\|_2) = \frac{1}{k} \sum_{i=1}^{k} \|A_i\|_2.$$

Proof. Note that

$$\begin{aligned}
G(A; \|\cdot\|_1, \|\cdot\|_2) &= \frac{E[\|Az\|_1]}{E[\|z\|_1]} \\
&= \frac{E\left[|A_1^{\mathsf{T}} z|\right] + \cdots + E\left[|A_k^{\mathsf{T}} z|\right]}{E[|z_1|] + \cdots + E[|z_k|]} \\
&= \frac{\|A_1\|_2 E[|z_1|] + \cdots + \|A_k\|_2 E[|z_1|]}{E[|z_1|] + \cdots + E[|z_k|]} \\
&= \frac{1}{k} \sum_{i=1}^{k} \|A_i\|_2,
\end{aligned}$$

in which the third equality follows from Lemma 2. $\blacksquare$

AN APPLICATION: ERROR MATRICES IN INPUT-OUTPUT ANALYSIS

Our interest in defining expected-value norms on matrices arose from recent research into the effect of aggregation in input-output modeling. In this section we will discuss the problems that arise from the use of induced norms in input-output aggregation, and how the expected-value norms solve these problems.

An input-output model is a linear economic model showing the relationship between "intermediate demand," "final demand," and output. Intermediate demand is the use of a commodity directly in the production of another commodity; final demand is the use of a commodity for any purpose (consumption, government spending, investment, and net exports). The fundamental assumption of input-output analysis is that intermediate inputs are a linear function of output.

Suppose that there are n commodities. Let R^n denote real n-dimensional space. Let $x \in R^n$ be output by commodity; let $y \in R^n$ be final demand by commodity; and let $A = [a_{ij}] \in M_n$ be the "direct requirements" matrix, defined so that a_{ij} is the amount of good i required to produce a single unit of good j.

In order for an economic system to be in equilibrium, there must be equality between the quantity of every commodity demanded and supplied. Thus equilibrium requires that output must equal the sum of intermediate demand and final demand for every commodity, so $x = Ax + y$, or

$$y = (I - A)x.$$

The above equation and its inverse, $x = (I - A)^{-1}y$, are the two fundamental equilibrium conditions used in input-output analysis.

Input-output analysis is done at various levels of aggregation; models range in size from several thousand commodities to fewer than ten. There is a large literature—almost exclusively theoretical—on the existence of aggregation error. Aggregation error is modeled in the following way. Consider a model with n *microcommodities* versus one in which the microcommodities have been aggregated into m *macrocommodities*, with $m < n$. For a given partition of the n microcommodities into a m macrocommodities, let $S \in M_{m,n}$ be a 0, 1 matrix, with $s_{ij} = 1$ if microcommodity i is to be included in macrocommodity j, and 0 otherwise. Let $A \in M_n$ and $B \in M_m$ be the n- and m-commodity direct requirements matrices. Observe the matrix S performs aggregation, mapping n-dimensional microcommodity vectors into m-dimensional macrocommodity vectors.

Suppose we are given the microcommodity output x and wish to compute the macro-commodity final demand from it. We could do this in either of two ways: by computing $S(I-A)x$, which uses the disaggregated model to compute the microcommodity final demand and then aggregates the result; or by computing $(I-B)Sx$, which computes the macrocommodity output and then uses the aggregated model to compute the final demand. Assuming that the disaggregated model is actually correct, the difference between these two is the aggregation error. That difference is the following:

$$S(I-A)x-(I-B)Sx=(BS-SA)x.$$

The dependence of the above expression on the precise vector x is unfortunate, so it is usual to use $|||\, BS-SA\, |||$, for some matrix norm $|||\cdot|||$. There are other matrices which norms are also used to measure aggregation error [see Howe and Johnson (1989a, b), which have extensive bibliographies of the aggregation literature]. The expression $|||\, BS-SA\, |||$, however, is sufficient for the point we wish to make.

Although there is a large theoretical literature on aggregation error (in which a number of prominent writers have called for aggregation error to be minimized), there has been little actual computational work done in error minimization. Three recent papers have appeared which deal computationally with the problem of minimizing aggregation error: Harrison and Manning (1987), Howe and Stabler (1989), and Howe and Johnson (1989b).

The introduction above gave two motivations for expected-value norms; both had to do with aspects of induced norms that are unfortunate for some applications. One aspect is that since an induced norm of a matrix records the extreme value by which a matrix stretches a vector, it may be unrepresentative of the amount that a vector is typically stretched. Moreover, in applications where the norm, $|||\, A\, |||$, of a matrix is being minimized, minimization of $|||\, A\, |||$ may be misleading in comparison with minimizing a more representative value of $\|Ax\|$. Both of these drawbacks are illustrated in Howe and Stabler (1989). That paper reports on aggregations of a multiregional input-output model of Canada that minimize $|||\, BS-SA\, |||$. The paper reports on the results of minimizing aggregation error for the reduction of the original matrix 176 microcommodities down to as few 34 macrocommodities. The norm used is the (induced) maximum-absolute-column-sum norm, $|||\cdot|||_1$. The results that are reported are difficult to interpret and tend to be erratic—highly dependent on the precise specification of the problem.

Difficulties with induced norms lead Howe and Stabler to adopt a simple expected-value norm $E[\|(BS-SA)x\|]$, where x is a random variable with a specified distribution. The expected-value approach is then shown to yield results that are readily interpretable and less erratic. A similar expected-value approach was adopted by Harrison and Manning.

It is not desirable, however, to abandon the use of matrix norms, since matrix norms are useful in the *theory* of aggregation.

Thus the preceding sections have established that the nicer computational results associated with the expected-value approach can still be obtained within the context of a (generalized) matrix norm. Most of the theoretical value of the norm-based approach for input-output results from the use of the triangle inequality and submultiplicativity. The above class of expected-value norms has been shown (of course) to have triangle inequality; although it does not have submultiplicativity, the property given in Theorem 1 is almost as useful.

COMMENTS

We note two alternative definitions of G also suggested by the motivating discussion in the introduction. Again for norms n_1 and n_2 on C^k and $A \in M_k$, define

$$G'(A; n_1, n_2) = E\left[\frac{n_1(Az)}{n_1(z)}\right]$$

and

$$G''(A; n_1, n_2) = E\left[n_1(Az)\right].$$

In both cases, the expectation is taken relative to the uniform distribution on $B(n_2)$, as before. Each of these also gives a generalized matrix norm, but both seem less tractable analytically.

REFERENCES

Harrison, Glen W. and Manning, Richard. 1987. Best approximate aggregation of input-output systems, *J. Amer. Statist. Assoc.* 82:1027–1031.

Horn, Roger A. and Johnson, Charles R. 1985. *Matrix Analysis*, Cambridge U.P., Cambridge.

Howe, Eric C. and Johnson, Charles R. 1989a. Linear aggregation of input-output models, *SIAM J. Matrix Anal. Appl.* 10:65–79.

——. 1989b. Analytic problems in the aggregation of input-output models, in *Applications of Matrix Theory*, Oxford U.P., Oxford.

Howe, Eric C. and Stabler, Jack C. 1989. Canada divided: The optimal division of an economy into regions, *J. Regional Sci.* 29:191–211.

Received 3 March 1989; final manuscript accepted 1 June 1989